

# SEQUENCE LISTING

<110> Tang, Jordan J.N.

Hong, Lin

Ghosh, Arun K.

<120> Inhibitors of Memapsin 2 and Use Thereof

<130> OMRF 182

<140> Not Yet Assigned

<141> 2000-06-27

<150> 60/141,363

<151> 1999-06-28

<150> 60/168,060

<151> 1999-11-30

<150> 60/177,836

<151> 2000-01-25

<150> 60/178,368

<151> 2000-01-27

<150> 60/210,292

<151> 2000-06-08

<160> 31

<170> PatentIn Ver. 2.1

<210> 1

<211> 3252

<212> DNA

<213> Homo sapiens

<400> 1

```
gcgggagtg tgcctgcca cggcaccag cagggcatcc ggctgcccct gcgcagcggc 60
ctggggggcg ccccccctggg gctgcggtcg ccccgaggaga ccgacgaaga gcccaggag 120
cccgcccgga ggggcagctt tgtggagatg gtggacaacc tgaggggcaa gtcggggcag 180
ggctactacg tggagatgac cgtgggcagc ccccgcgaga cgctcaacat cctgggtgat 240
acaggcagca gtaactttgc agtgggtgct gccccccacc ccttcttgca tcgctactac 300
cagaggcagc tgtccagcac atacggggac ctccggaagg gtgtgtatgt gccctacacc 360
cagggcaagt ggggaaggga gctgggcacc gacctggtaa gcatccccaac tggccccaac 420
gtcactgtgc gtgccaacat tctgcccac actgaatcag acaagttctt catcaacggc 480
tccaactggg aaggcatcct ggggctggcc tatgctgaga ttgccaggcc tgacgactcc 540
ctggagcctt tctttgactc tctggtaaag cagacccacg ttcccaacct cttctccctg 600
```

```

cagctttgtg gtgctggcct cccctcaac cagtctgaag tgctggcctc tgcggaggg 660
agcatgatca ttggagggtat cgaccactcg ctgtacacag gcagtctctg gtatacacc 720
atccggcggg agtgggtatta tgagggtgatc attgtgcggg tggagatcaa tggacaggat 780
ctgaaaatgg actgcaagga gtacaactat gacaagagca ttgtggacag tggcaccacc 840
aaccttcgtt tgcccaagaa agtgtttgaa gctgcagtca aatccatcaa ggcagcctcc 900
tccacggaga agttccctga tggtttcttg ctaggagagc agctgggtgtg ctggcaagca 960
ggcaccacc cttggaacat tttcccagtc atctcactct acctaattggg tgagggttacc 1020
aaccagtcc tccgcatcac catccttccg cagcaatacc tgcggccagt ggaagatgtg 1080
gccacgtccc aagacgactg ttacaagttt gccatctcac agtcatccac gggcactgtt 1140
atgggagctg ttatcatgga gggcttctac gttgtctttg atcggggccg aaaacgaatt 1200
ggctttgtcg tcagcgcttg ccatgtgcac gatgagttca ggacggcagc ggtggaaggc 1260
ccttttgtca ccttgacat ggaagactgt ggctacaaca tccacagac agatgagtca 1320
accctcatga ccatagccta tgtcatggct gccatctgcg cctcttcat gctgccactc 1380
tgcctcatgg tgtgtcagtg gcgctgcctc cgctgcctgc gccagcagca tgatgacttt 1440
gctgatgaca tctccctgct gaagtgaaga ggcccatggg cagaagatag agattcccct 1500
ggaccacacc tccgtggttc actttggtca caagtaggag acacagatgg cacctgtggc 1560
cagagcacct caggaccctc cccaccacc aaatgcctct gccttgatgg agaaggaaaa 1620
ggctggcaag gtgggttcca gggactgtac ctgtaggaaa cagaaaagag aagaaagaag 1680
cactctgctg gcgggaatac tcttggtcac ctcaaattta agtcgggaaa ttctgctgct 1740
tgaaacttca gccctgaacc tttgtccacc attcctttaa attctccaac ccaaagtatt 1800
cttcttttct tagtttcaga agtactggca tcacacgcag gttaccttgg cgtgtgtccc 1860
tgtggtaccc tggcagagaa gagaccaagc ttgtttccct gctggccaaa gtcagtagga 1920
gaggatgcac agtttgctat ttgctttaga gacagggact gtataaacia gcctaacatt 1980
ggtgcaaaga ttgcctcttg aattaaaaaa aaactagatt gactatttat acaaatgggg 2040
gcggttgga agaggagaag gagaggaggt acaaagacag ggaatagtgg gatcaaagct 2100
aggaaaggca gaaacacaac cactcaccag tcctagtttt agacctcctc tccaagatag 2160
catcccatct cagaagatgg gtgttggttt caatgttttc ttttctgtgg ttgcagcctg 2220
accaaagtg agatgggaag ggcttatcta gccaaagagc tcttttttag ctctcttaaa 2280
tgaagtgcc actaagaagt tccacttaac acatgaattt ctgccatatt aatttcattg 2340
tctctatctg aaccaccctt tattctacat atgataggca gcactgaaat atcctaacc 2400
cctaagctcc aggtgcctg tgggagagca actggactat agcagggctg ggctctgtct 2460
tcttggtcat aggtcactc tttccccc aaacttctct ggagctttgc agccaagggtg 2520
ctaaaaggaa taggtaggag acctcttcta tctaactcct aaaagcataa tgttgacat 2580
tcattcaaca gctgatgcc tataaccct gcctggattt cttcctatta ggctataaga 2640
agtagcaaga tctttacata attcagagt gtttcattgc cttcctacc tctctaattg 2700
ccctccatt tatttgacta aagcaterca cagtggcact agcattatac caagagtatg 2760
agaaatacag tgctttatgg ctctaaccatt actgccttca gtatcaaggc tgctggaga 2820
aaggatggca gcctcagggc ttccttatgt cctccaccac aagagctcct tgatgaagg 2880
catcttttcc cctatcctg ttcttccct cccgcctcct aatgggtacgt gggtaaccag 2940
gctgggtctt gggctaggta gtggggacca agttcattac ctccctatca gttctagcat 3000
agtaaactac ggtaccagt ttagtgggaa gagctgggtt ttcctagtat acccactgca 3060
tcctactcct acctggtcaa cccgctgctt ccaggatagg gacctgctaa gtgtggaatt 3120
acctgataag ggagaggga atacaaggag ggcctctggt gttcctggcc tcagccagct 3180
gccmcaagc cataaaccaa taaamcaaga atactgagtc taaaaaaaaa aaaaaaaaaa 3240
aaaaaaaaaa aa 3252

```

<211> 488

<212> PRT

<213> Homo sapiens

<220>

<223> Purified Memapsin 2

<220>

<223> Amino Acids 28-48 are remnant putative propeptide residues

<220>

<223> Amino Acids 58-61, 78, 80, 82-83, 116, 118-121, 156, 166, 174, 246, 274, 276, 278-281, 283, and 376-377 are residues in contact with the OM99-2 inhibitor

<220>

<223> Amino acids 54-57, 61-68, 73-80, 86-89, 109-111, 113-118, 123-134, 143-154, 165-168, 198-202, and 220-224 are N-lobe Beta Strands

<220>

<223> Amino Acids 184-191 and 210-217 are N-lobe Helices

<220>

<223> Amino acids 237-240, 247-249, 251-256, 259-260, 273-275, 282-285, 316-318, 331-336, 342-348, 354-357, 366-370, 372-375, 380-383, 390-395, 400-405, and 418-420 are C-lobe Beta Strands

<220>

<223> Amino Acids 286-299, 307-310, 350-353, 384-387, and 427-431 are C-lobe Helices

<400> 2

Ala Gly Val Leu Pro Ala His Gly Thr Gln His Gly Ile Arg Leu Pro  
1 5 10 15

Leu Arg Ser Gly Leu Gly Gly Ala Pro Leu Gly Leu Arg Leu Pro Arg  
20 25 30

Glu Thr Asp Glu Glu Pro Glu Glu Pro Gly Arg Arg Gly Ser Phe Val  
35 40 45

Glu Met Val Asp Asn Leu Arg Gly Lys Ser Gly Gln Gly Tyr Tyr Val  
50 55 60

Glu Met Thr Val Gly Ser Pro Pro Gln Thr Leu Asn Ile Leu Val Asp			
65	70	75	80
Thr Gly Ser Ser Asn Phe Ala Val Gly Ala Ala Pro His Pro Phe Leu			
	85	90	95
His Arg Tyr Tyr Gln Arg Gln Leu Ser Ser Thr Tyr Arg Asp Leu Arg			
	100	105	110
Lys Gly Val Tyr Val Pro Tyr Thr Gln Gly Lys Trp Glu Gly Glu Leu			
	115	120	125
Gly Thr Asp Leu Val Ser Ile Pro His Gly Pro Asn Val Thr Val Arg			
	130	135	140
Ala Asn Ile Ala Ala Ile Thr Glu Ser Asp Lys Phe Phe Ile Asn Gly			
145	150	155	160
Ser Asn Trp Glu Gly Ile Leu Gly Leu Ala Tyr Ala Glu Ile Ala Arg			
	165	170	175
Pro Asp Asp Ser Leu Glu Pro Phe Phe Asp Ser Leu Val Lys Gln Thr			
	180	185	190
His Val Pro Asn Leu Phe Ser Leu Gln Leu Cys Gly Ala Gly Phe Pro			
	195	200	205
Leu Asn Gln Ser Glu Val Leu Ala Ser Val Gly Gly Ser Met Ile Ile			
	210	215	220
Gly Gly Ile Asp His Ser Leu Tyr Thr Gly Ser Leu Trp Tyr Thr Pro			
225	230	235	240
Ile Arg Arg Glu Trp Tyr Tyr Glu Val Ile Ile Val Arg Val Glu Ile			
	245	250	255
Asn Gly Gln Asp Leu Lys Met Asp Cys Lys Glu Tyr Asn Tyr Asp Lys			
	260	265	270
Ser Ile Val Asp Ser Gly Thr Thr Asn Leu Arg Leu Pro Lys Lys Val			
	275	280	285
Phe Glu Ala Ala Val Lys Ser Ile Lys Ala Ala Ser Ser Thr Glu Lys			
	290	295	300
Phe Pro Asp Gly Phe Trp Leu Gly Glu Gln Leu Val Cys Trp Gln Ala			
305	310	315	320

Gly Thr Thr Pro Trp Asn Ile Phe Pro Val Ile Ser Leu Tyr Leu Met  
 325 330 335

Gly Glu Val Thr Asn Gln Ser Phe Arg Ile Thr Ile Leu Pro Gln Gln  
 340 345 350

Tyr Leu Arg Pro Val Glu Asp Val Ala Thr Ser Gln Asp Asp Cys Tyr  
 355 360 365

Lys Phe Ala Ile Ser Gln Ser Ser Thr Gly Thr Val Met Gly Ala Val  
 370 375 380

Ile Met Glu Gly Phe Tyr Val Val Phe Asp Arg Ala Arg Lys Arg Ile  
 385 390 395 400

Gly Phe Ala Val Ser Ala Cys His Val His Asp Glu Phe Arg Thr Ala  
 405 410 415

Ala Val Glu Gly Pro Phe Val Thr Leu Asp Met Glu Asp Cys Gly Tyr  
 420 425 430

Asn Ile Pro Gln Thr Asp Glu Ser Thr Leu Met Thr Ile Ala Tyr Val  
 435 440 445

Met Ala Ala Ile Cys Ala Leu Phe Met Leu Pro Leu Cys Leu Met Val  
 450 455 460

Cys Gln Trp Arg Cys Leu Arg Cys Leu Arg Gln Gln His Asp Asp Phe  
 465 470 475 480

Ala Asp Asp Ile Ser Leu Leu Lys  
 485

<210> 3  
 <211> 503  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <223> Pro-memapsin 2

<220>  
 <223> Amino Acids 1-15 are vector-derived residues

<220>  
 <223> Amino Acids 16-64 are a putative pro peptide

<220>

<223> Amino Acids 1-13 are the T7 promoter

<220>

<223> Amino Acids 16-456 are Pro-memapsin 2-T1

<220>

<223> Amino Acids 16-421 are Promemapsin 2-T2

<400> 3

Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg Gly Ser Met Ala

1

5

10

15

Gly Val Leu Pro Ala His Gly Thr Gln His Gly Ile Arg Leu Pro Leu

20

25

30

Arg Ser Gly Leu Gly Gly Ala Pro Leu Gly Leu Arg Leu Pro Arg Glu

35

40

45

Thr Asp Glu Glu Pro Glu Glu Pro Gly Arg Arg Gly Ser Phe Val Glu

50

55

60

Met Val Asp Asn Leu Arg Gly Lys Ser Gly Gln Gly Tyr Tyr Val Glu

65

70

75

80

Met Thr Val Gly Ser Pro Pro Gln Thr Leu Asn Ile Leu Val Asp Thr

85

90

95

Gly Ser Ser Asn Phe Ala Val Gly Ala Ala Pro His Pro Phe Leu His

100

105

110

Arg Tyr Tyr Gln Arg Gln Leu Ser Ser Thr Tyr Arg Asp Leu Arg Lys

115

120

125

Gly Val Tyr Val Pro Tyr Thr Gln Gly Lys Trp Glu Gly Glu Leu Gly

130

135

140

Thr Asp Leu Val Ser Ile Pro His Gly Pro Asn Val Thr Val Arg Ala

145

150

155

160

Asn Ile Ala Ala Ile Thr Glu Ser Asp Lys Phe Phe Ile Asn Gly Ser

165

170

175

Asn Trp Glu Gly Ile Leu Gly Leu Ala Tyr Ala Glu Ile Ala Arg Pro

180

185

190

Asp Asp Ser Leu Glu Pro Phe Phe Asp Ser Leu Val Lys Gln Thr His

195

200

205







[illegible]

5

**<211> 12**

<212> PRT

### <213> Artificial Sequence

**<220>**

<223> Description of Artificial Sequence: Synthetic Peptide

**<400> 7**

Lys Gly Gly Val Val Ile Ala Thr Val Ile Val Lys

1

5

10

<210> 8

<211> 4

<212> PRT

<213> Homo sapiens

**<400> 8**

Asp Thr Ser Gly

1

<210> 9

<211> 8

<212> PRT

<213> Homo sapiens

**<400> 9**

Leu Val Asn Met Ala Glu Gly Asp

1

5

<210> 10

<211> 28

**<212> DNA**

<213> Artificial Sequence

**<220>**

<223> Description of Artificial Sequence: Primer

**<400> 10**

ggtaagcatc ccccatggcc ccaacgtc

<210> 11  
<211> 28  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 11  
gacgttgggg ccatggggga tgcttacc

28

<210> 12  
<211> 34  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 12  
acgttgtctt tgatcgggcc cgaaaacgaa ttgg

34

<210> 13  
<211> 33  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 13  
ccaattcgtt ttcgggccc atcaaagaca acg

33

<210> 14  
<211> 27  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 14  
ccatccta at acgactcact atagggc

27

<210> 15  
<211> 23  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 15  
actcactata gggctcgagc ggc 23

<210> 16  
<211> 26  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 16  
cttttgagca agttcagcct ggttaa 26

<210> 17  
<211> 31  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 17  
gaggtggctt atgagtattt cttccagggt a 31

<210> 18  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 18  
tggcgacgac tcctggagcc cg 22

<210> 19  
<211> 24  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 19  
tgacaccaga ccaactggta atgg

24

<210> 20  
<211> 27  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 20  
catatggcgg gagtgctgcc tgcccac

27

<210> 21  
<211> 38  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 21  
ggatcctcac ttcagcaggg agatgtcatc agcaaagt

38

<210> 22  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Oxidized  
Insulin B-chain

<220>

**090768**

His Leu Xaa Gly Ser His Leu Val

5

<211> 8

### <213> Artificial Sequence

<223> Description of Artificial Sequence: Oxidized Insulin B-chain

<223> Xaa at site 1 represents cysteic acid

Xaa Gly Glu Arg Gly Phe Phe Tyr

5

**<211> 5**

### <213> Artificial Sequence

<223> Description of Artificial Sequence: Synthetic Peptide

Val Gly Ser Gly Val

5

<211> 7

<213> Artificial Sequence

<223> Description of Artificial Sequence: Synthetic Peptide

13

Val Gly Ser Gly Val Leu Leu  
1 5

<210> 26

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic  
Peptide

<400> 26

Gly Val Leu Leu Ser Arg Lys  
1 5

<210> 27

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Inhibitors

<400> 27

Val Asn Leu Ala Ala Glu Phe  
1 5

<210> 28

<211> 8

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Inhibitors

<400> 28

Glu Val Asn Leu Ala Ala Glu Phe  
1 5

<210> 29

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic  
Peptide

<400> 29

Asn Leu Ala Ala

1

<210> 30

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic  
Peptide

<400> 30

Val Gly Ser Gly Val Leu Leu Ser Arg Lys

1

5

10

<210> 31

<211> 326

<212> PRT

<213> Homo sapiens

<220>

<223> Amino acids 2-5, 6-9, 13-20, 25-32, 65-67, 69-74,  
79-87, 89-91, 99-106, 119-122, 150-154, 164-167,  
180-183, 191-194, 196-199, 201-204, 210-214,  
221-223, 258-262, 265-269, and 275-278 are Beta  
Strands

<220>

<223> Amino acids 281-284, 286-288, 298-301, 310-315,  
and 319-324 are Beta strands

<220>

<223> Amino acids 48-51, 111-114, 136-142, 225-234,  
249-254, 271-274, and 303-306 are Helices

<220>

<223> Amino acids 12-13, 30, 32, 34-35, 73-77, 111, 117,  
120, 189, 213, 215, 217-220, 287, 289, 291, 298,  
and 300 are residues in contact with pepstatin.





